



---

# The Future of Scientific Computing

*Horst D. Simon*

NERSC, Division Director

May 6, 2000





***"Technology does not drive change at all.  
Technology merely enables change. It's our  
collective cultural response to the options and  
opportunities presented by technology that drives  
change."***

**Paul Saffo  
Institute for the Future  
Menlo Park, California**



***"It's hard to make predictions, especially about the future."***

**Yogi Berra**

# Overview

---



- **Retrospective: changes in the 1990s**
- **Extrapolation to the near future up to 2010**
- **The end of Moore's Law in about 2020**
- **Beyond 2025**

# Things that did **not** happen in the last five years

---



**1992 predictions (after Forest Baskett, SGI):**

- **TV and PC converge**
- **interactive TV**
- **video servers instead of video stores**
- **Apple/IBM/Motorola**
- **Intel makes a mistake**
- **MPPs go mainstream**

# 1990s: Technology

---



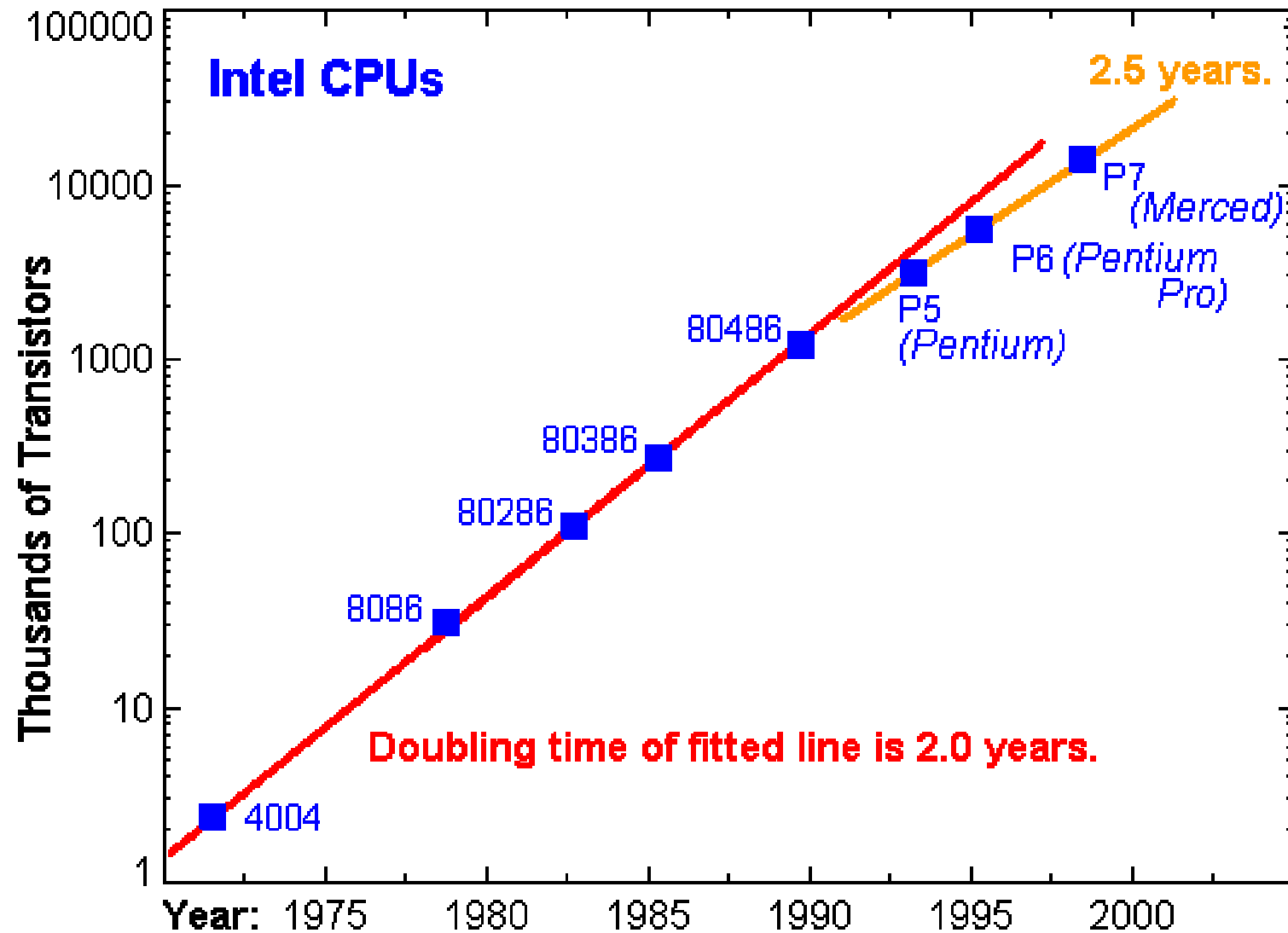
**In the 1980's there have been fundamental changes in the microprocessor development ("killer micros")**

- dramatic increase in number of transistors available per chip**
- architectural advances including the use of RISC ideas, pipelining and caches**
- as a result CPU performance has improved by a factor of 1.5 to 2.0 per year**

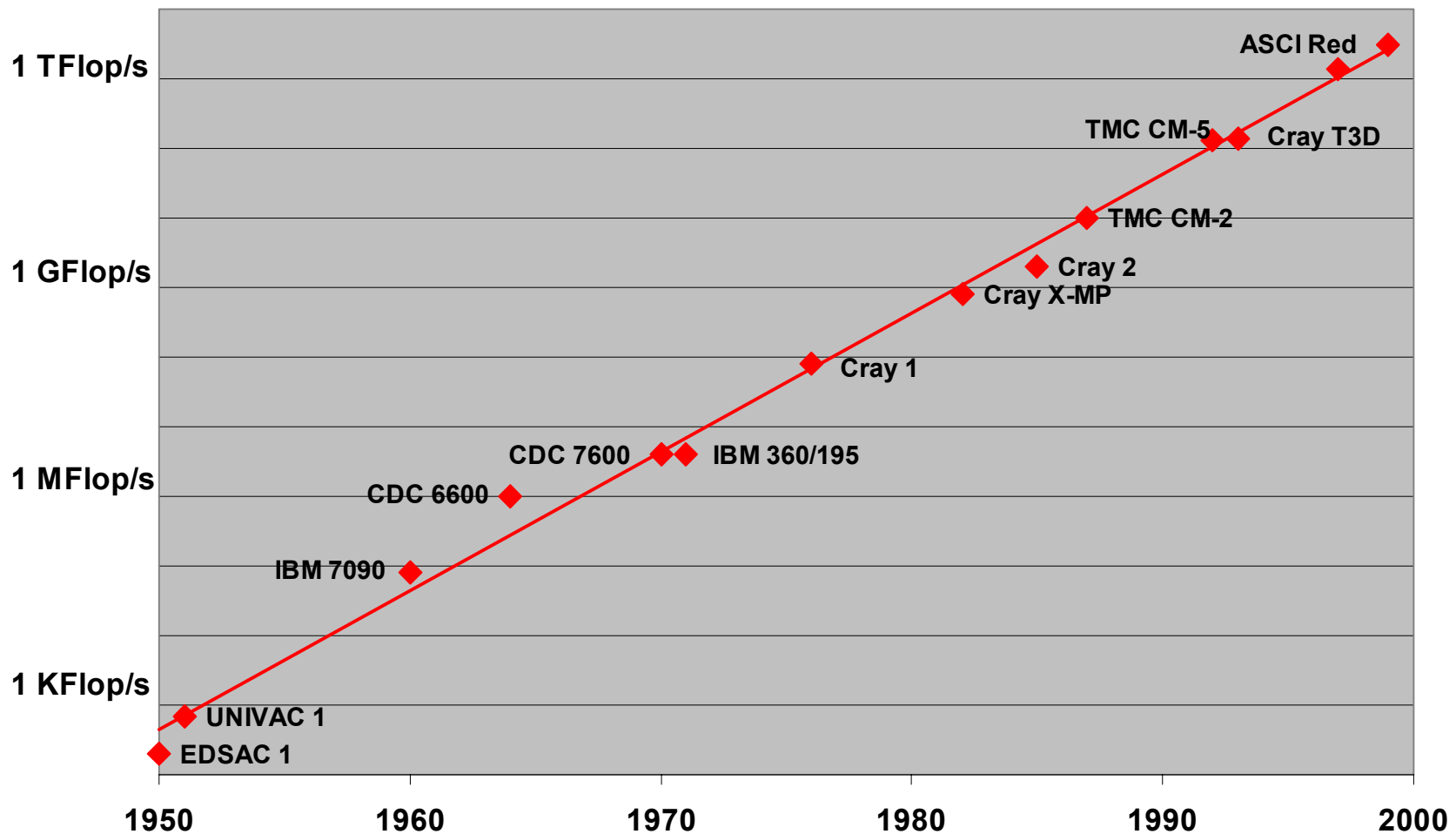
**Maturation in the late 80s**

**Full impact in the early 90s**

# Moore's Law



# Impact of Moore's Law on HPC





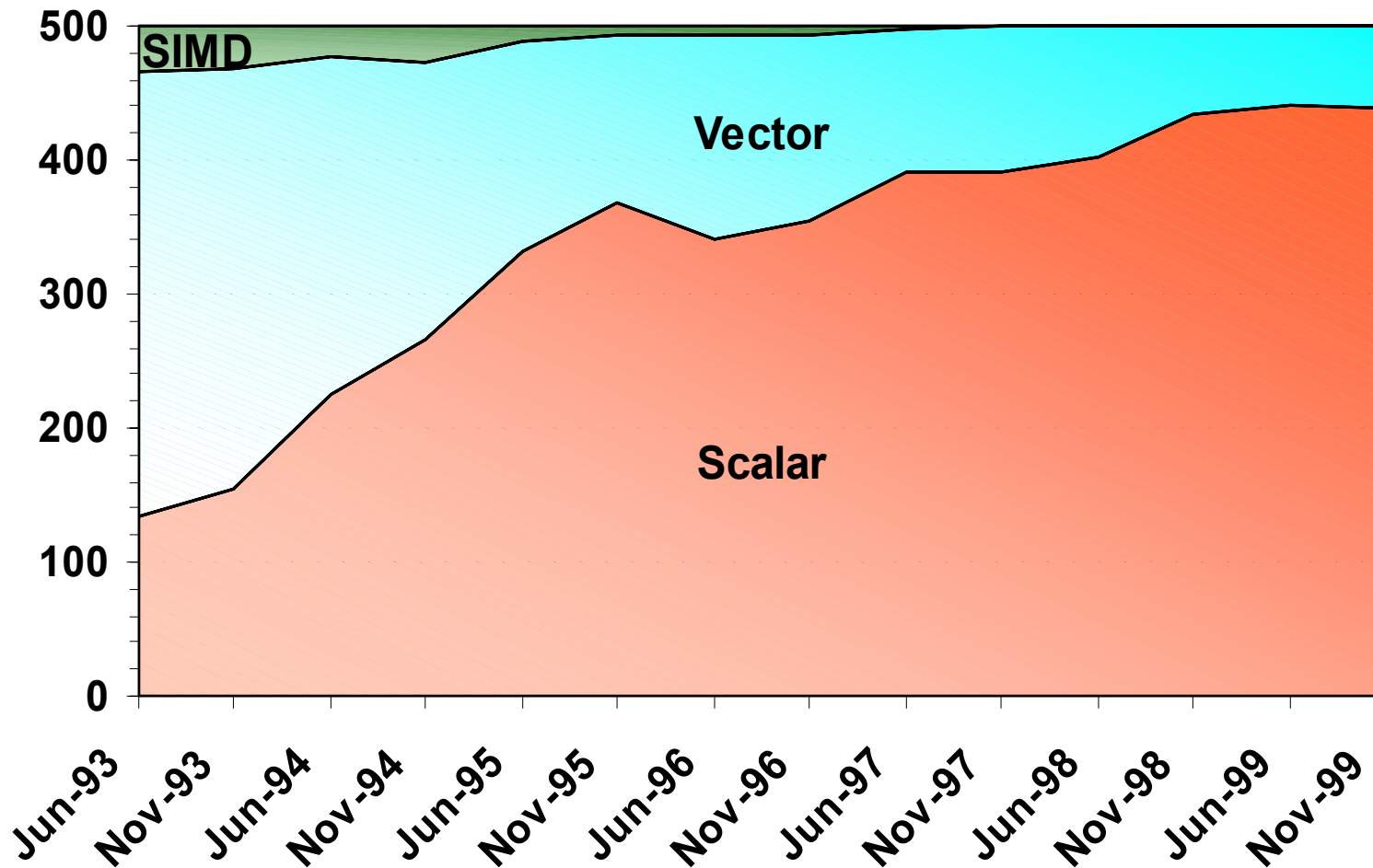
# TOP500 List

---



- Published twice a year with the 500 most powerful supercomputers in actual use
- Ranked according to LINPACK R<sub>max</sub>
- Data available since 1993
- For details see <http://www.top500.org/>

# Processor Design as Seen in the TOP500



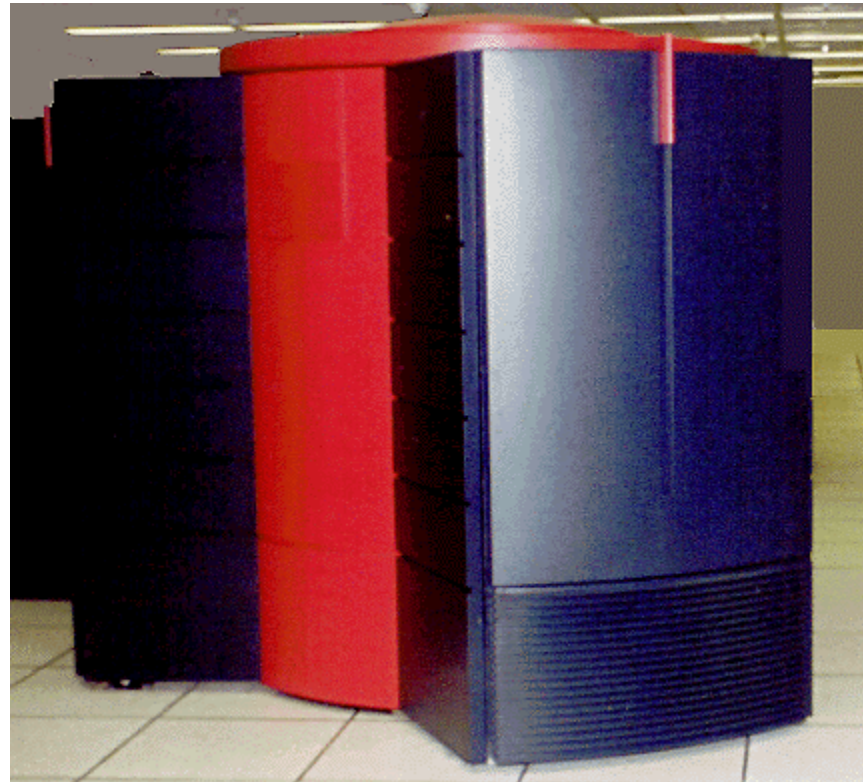
# NERSC-1

## Cray C90 installed in Dec. 1991

---



- Cray C90 installed in December 1991
- ended contract with CCC for a Cray-3
- stable high end production platform for seven years until 12/31/98



# NERSC- 2

## Cray T3E-900 installed in 1996

---



The 644 processor T3E-900 is one of the most powerful unclassified supercomputers in the U.S.

- eight out of twelve DOE Grand Challenge Projects compute at NERSC
- 50% of the resource dedicated to GC projects
- about 100 other projects allocated on the NERSC T3E-900
- 1997 GAO report judged NERSC to have the best MPP utilization (75%) -- 1999 utilization >90%



# NERSC-3

## IBM SP3 installed in 6/99

---



- New contract with IBM announced in April 1999
- IBM was clearly the best value for the primary award
  - provides the best absolute performance
  - has lowest absolute cost
  - provides the best price performance
  - provides acceptable functionality
  - guarantees performance - low risk

# NERSC-3 Supercomputer



- IBM selected to provide NERSC-3 (IBM SP3/RS 6000)
- Phase I: June 1999 installation
  - 608 processors
  - 410 gigaflop peak performance
  - Provides one teraflop NERSC capability
- Phase II: December 2000 comp
  - 2,432 processors
  - 3.2 teraflop peak performance
  - 4 teraflop total NERSC capability





# HPC Systems at NERSC in the 90s



	<b>NERSC-1</b> Cray C90	<b>NERSC-2</b> Cray T3E	<b>NERSC-3</b> IBM SP-3
Year of Installation	1991	1996	1999
Number of Processors	16	640	2048
Processor Technology	Custom ECL	Commodity CMOS	Commodity CMOS
Peak System Perform.	16 Gflop/s	580 Gflop/s	3000 Gflop/s
Architecture	Shared memory, parallel vector	Distributed memory	128 nodes with 16 processor SMP
System	Fully integrated custom system	Fully integrated custom system with commodity CPU and memory	Loosely integrated system with commodity system components
System Software	Vendor supplied, ready on delivery	Vendor supplied, completed after nearly 3 years development	Vendor supplied, contractual complete in about 3 years
Floor space	588 ft	360 ft	4000 ft
Power consumption	500 kW	288 kW	1400 kW

# Impact of Technology Transitions



	1994 – 1996 transition	1998 – 2000 transition
Economic Driver	Price performance of commodity processors and memory	16 – 64 CPU “sweet spot” for SMP technology in the commercial market place
Advantages of transition	Higher performance and better price performance	Higher performance
Challenges of transition	<ol style="list-style-type: none"><li>1) Applications transition to distributed memory, message passing model (MPI)</li><li>2) More complex system software (scheduling, checkpoint restarting)</li></ol>	<ol style="list-style-type: none"><li>1) Applications transition to hierarchical, distributed memory model (threads + MPI)</li><li>2) New development efforts for even more complex systems software</li><li>3) Increased cost of facilities</li></ol>

**Table 2.** Impact of the two technology transitions of the 1990s.



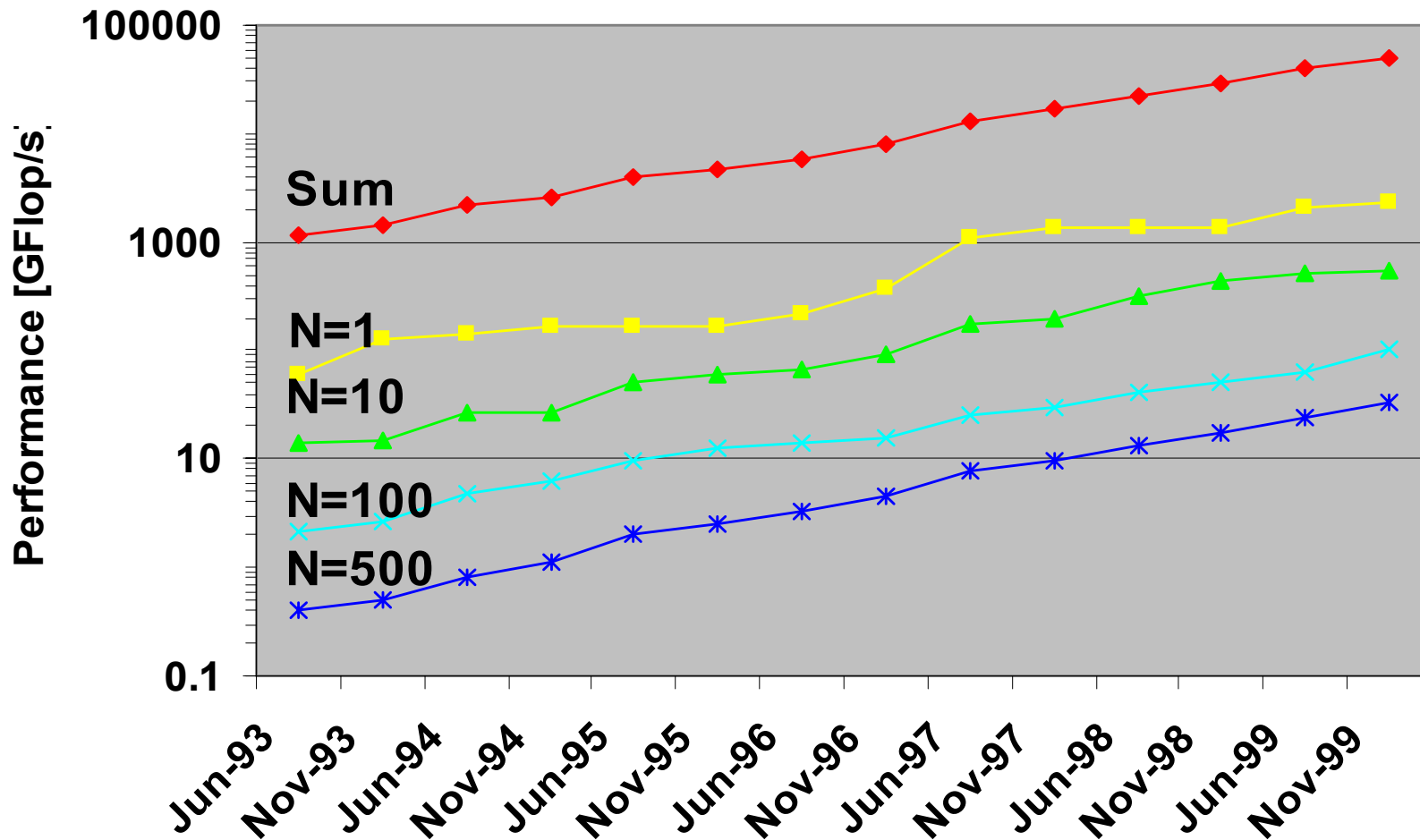
# Three Challenges

---



- applications that can tolerate an increase in communication latency and parallelism as well as a distributed, hierarchical memory model need to be written
- system software for increasingly complex, more difficult to manage, one-of-a-kind systems will have to be developed anew
- center management will be forced to take creative new approaches to solve the space and power requirements for the new systems.

# Performance Increases in the TOP500



# Analysis of TOP500 Data

---



- Annual performance growth about a factor of 1.82
- Two factors contribute almost equally to the annual total performance growth
- Processor number grows per year on the average by a factor of 1.30 and the
- Processor performance grows by 1.40 compared 1.58 of Moore's Law.
- For more details see paper by Dongarra, Meuer, Simon, and Strohmaier in Parallel Computing (to appear)

# The Revolution of 1994 - Major HPC Market Realignment

---



- 1991 Newcomers with CMOS and MPP technology (Intel, TMC, KSR) gain mind share and market share
- 1993 Cray, IBM, Convex go CMOS (T3D, SP 1/2, SPP 1000)
- 1994 TMC, KSR go out of business; SGI's SMP success
- 1995 HP buys Convex; Fujitsu, NEC introduce CMOS vector machines
- 1996 SGI buys Cray
- 1997 TOP500 take over by CMOS complete
- 2000 Tera buys Cray Division from SGI and renames itself Cray Inc.

# The Dead Supercomputer Society



- See <http://www.paralogos.com/DeadSuper/>
- list of 42 dead companies or projects from 1975 - today

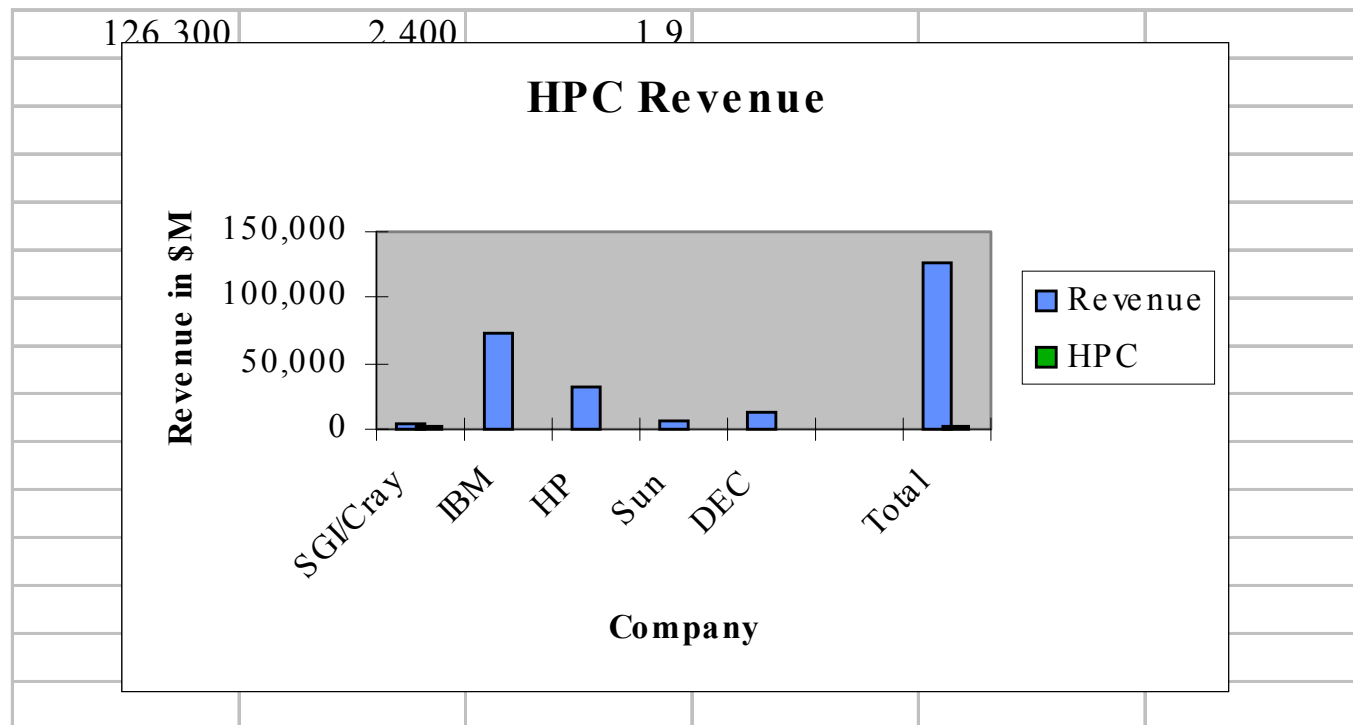


# Since 1997: The New HPC Marketplace



All major US HPC companies are now vertically integrated (SGI, IBM, HP, Sun, Compaq), with exception of Cray.

Almost all high-end products are based on workstation technology.



# 1997: The New HPC Marketplace

---



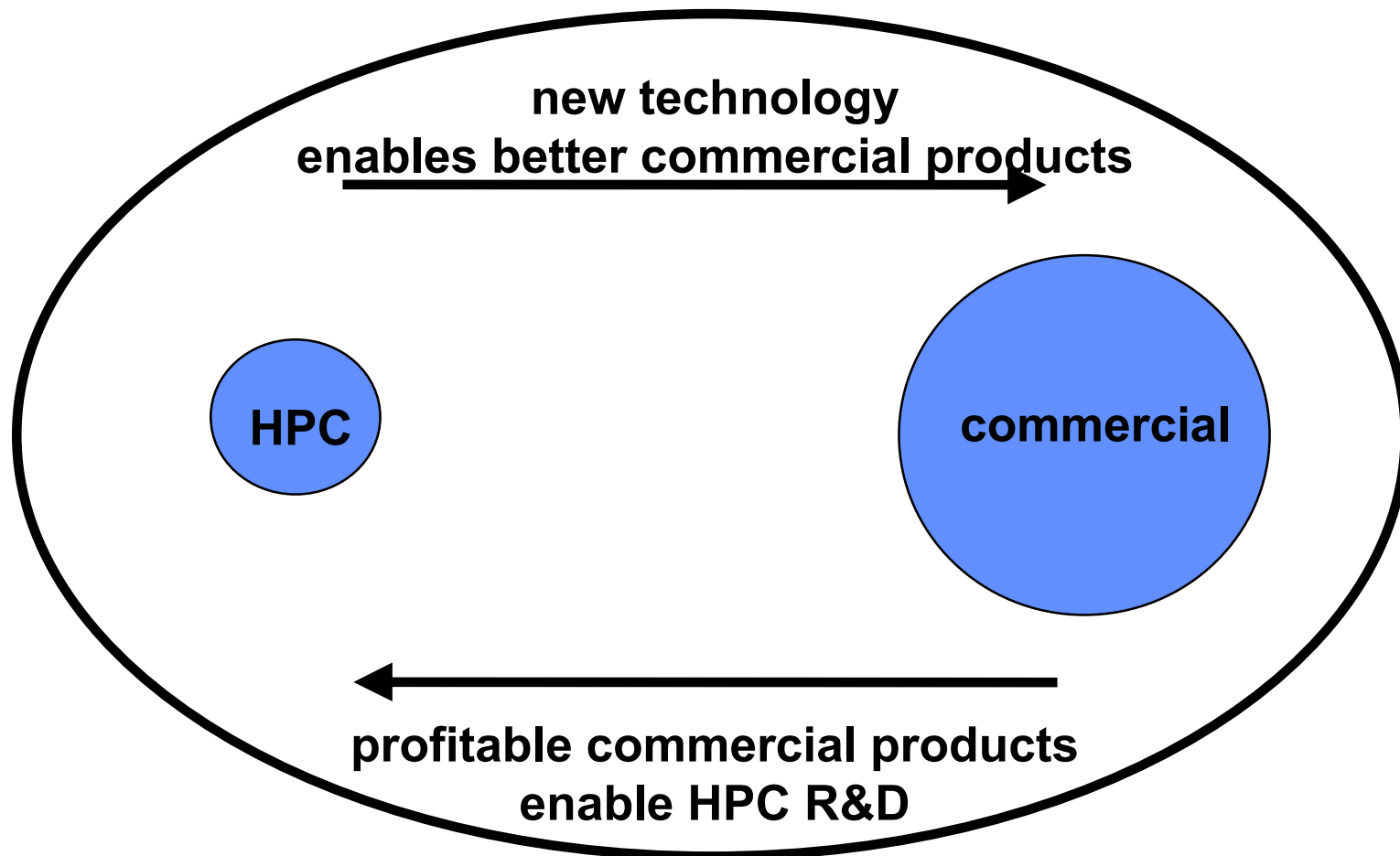
All these companies are in the computer **business**.

HPC customers must get used to a new role: they are no longer the center of attention.

Companies must have commitment to technology, and understand the potential of technology leverage from the high-end, in order to remain in the HPC business.

Fortunately for us, the HPC users, they all do understand that (for now).

# 1997: The HPC Business Model





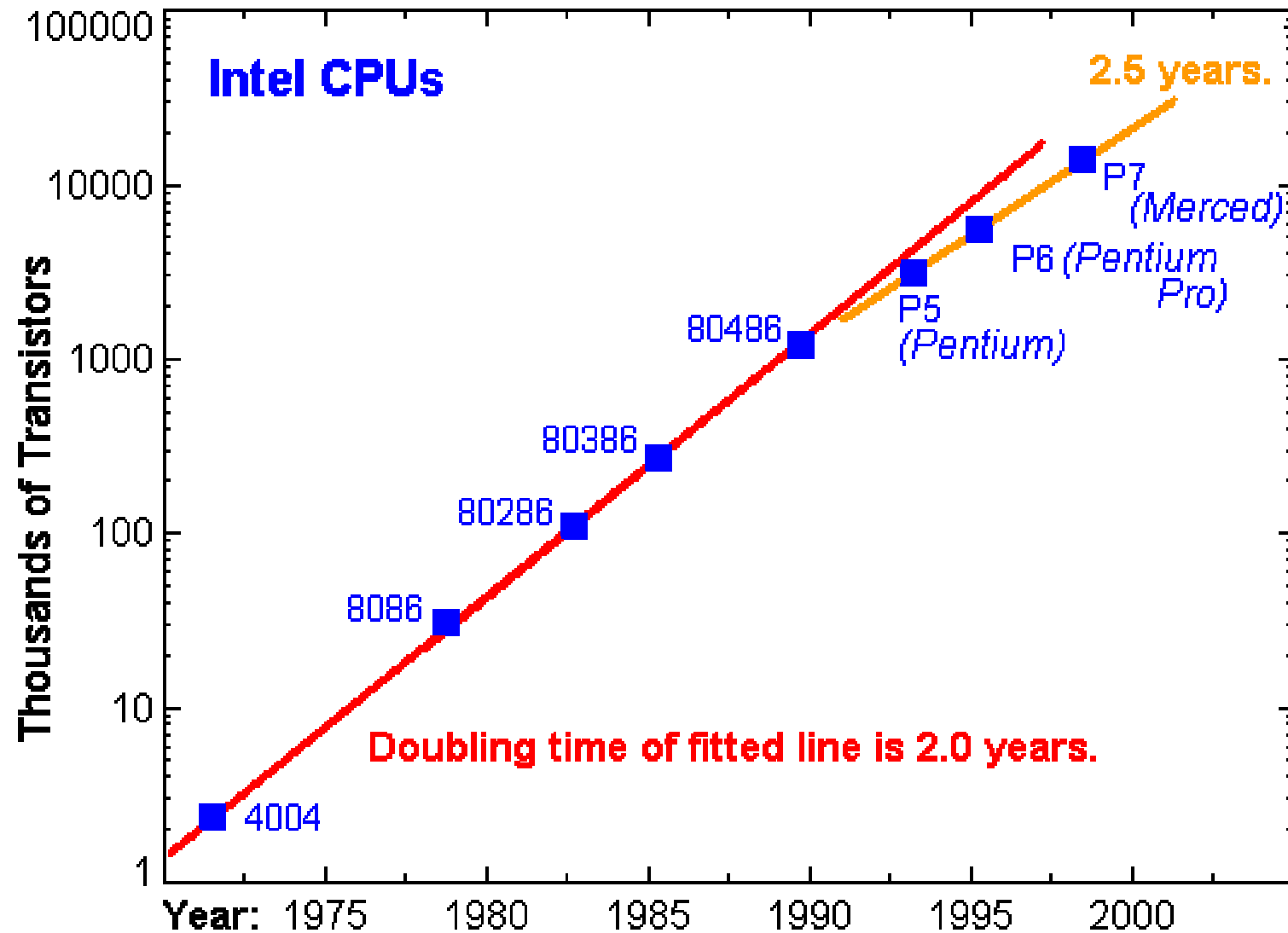
# Overview

---



- Retrospective: changes in the 1990s
- **Extrapolation to the near future up to 2010**
- The end of Moore's Law in about 2020
- Beyond 2025

# Moore's Law - the traditional (linear) view



# Moore's Wall - the real (exponential) view

---



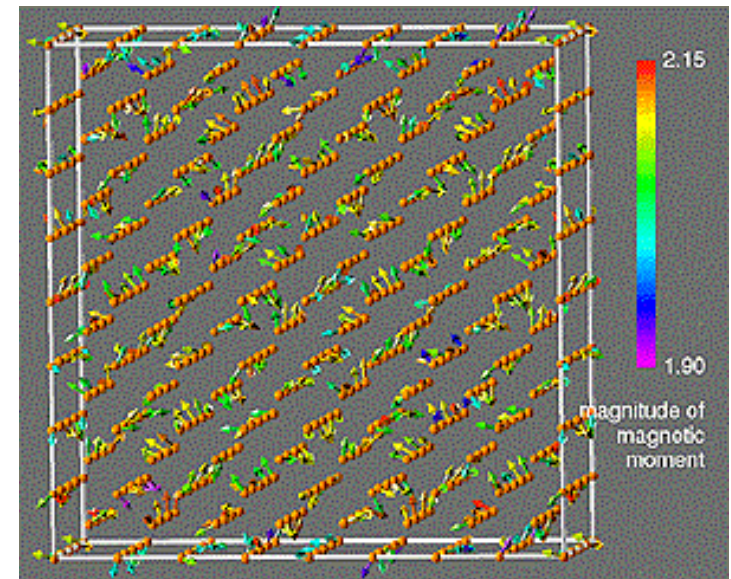
# Reality Check on Real Applications



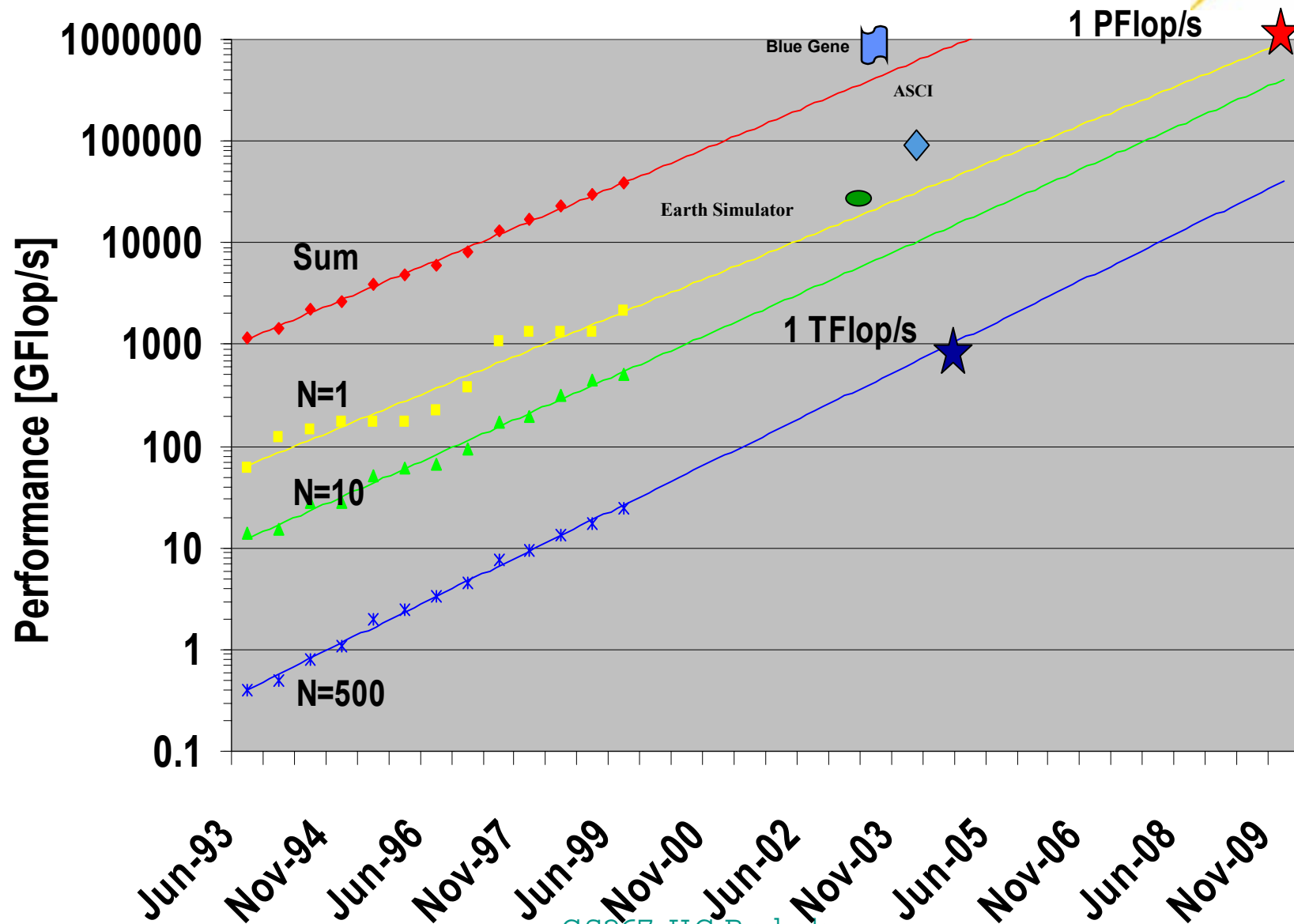
First complete application to break the 1Tflop/s sustained barrier in 1998.

Collaborators from DOE's Grand Challenge on Materials, Methods, Microstructure, and Magnetism.

1024-atom first-principles simulation of metallic magnetism in iron



# Extrapolation to the Next Decade



# Analysis of TOP500 Extrapolation

---



Based on the extrapolation from these fits we predict:

- First 100~TFlop/s system by 2005
- About 1--2 years later than the ASCI path forward plans.
- No system smaller than 1~TFlop/s should be able to make the Top500
- First Petaflop system available around 2009
- Rapid changes in the technologies used in HPC systems, therefore a projection for the architecture/technology is difficult
- Continue to expect rapid cycles of re-definition.

# 2000 - 2005: Technology Options

---



- Clusters
  - SMP nodes, with custom interconnect
  - PCs, with commodity interconnect
  - vector nodes (in Japan)
- Custom built supercomputers
  - Cray SV-2
  - IBM Blue Gene
  - HTMT
- Other technology to influence HPC
  - IRAM/PIM
  - Computational and Data Grids

# What Will a 10 Tflop/s System Look Like?



- The first ones are already on order  
Lawrence Livermore National Laboratory in US  
NERSC will have a 3 Tflop/s system in 2000

- Systems are large clusters  
SMP nodes in US  
Vector nodes in Japan



- Programming model:  
OpenMP and/or vectors to maximize node speed  
MPI for global communication



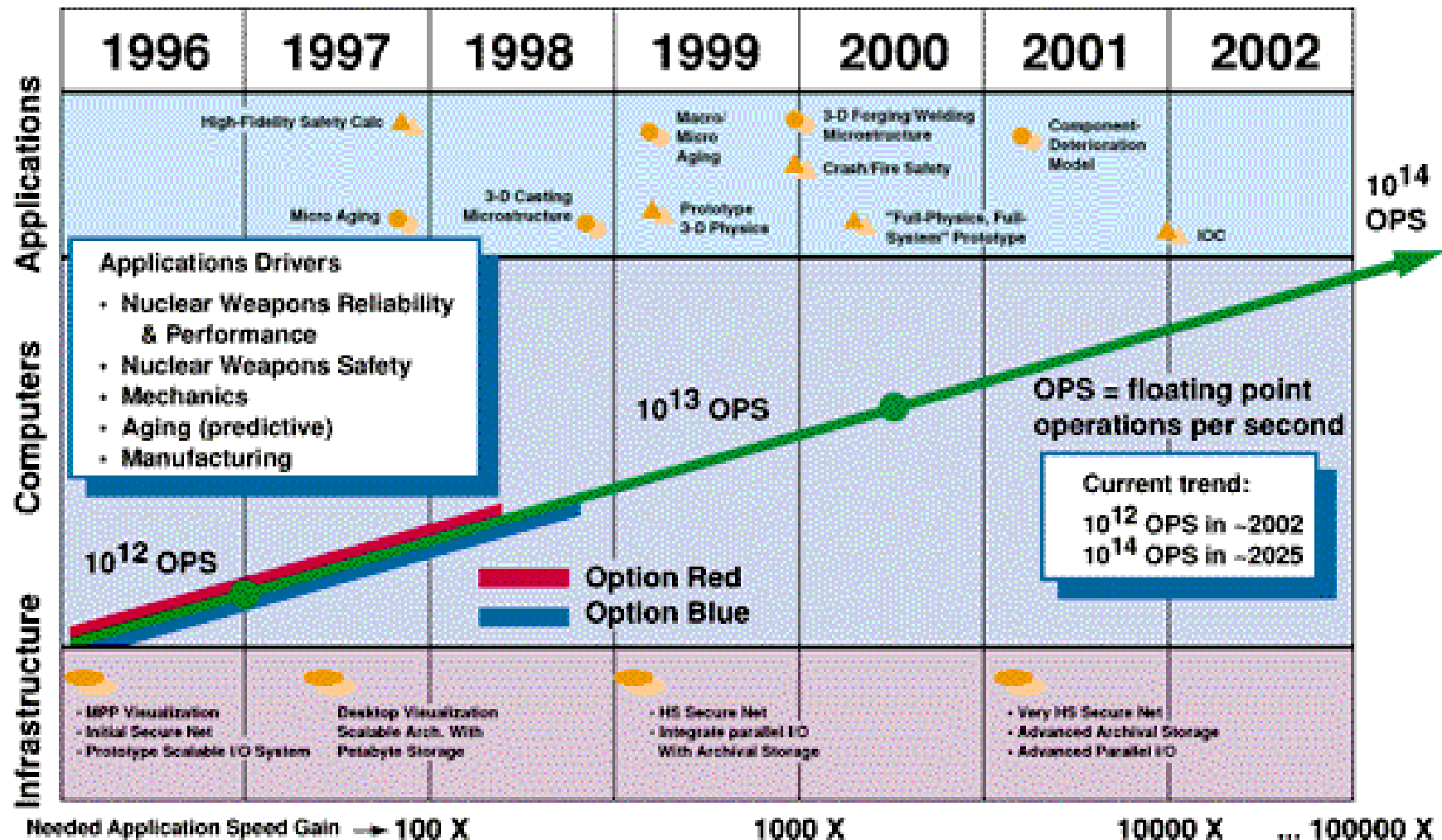


## ASCI - Accelerated Strategic Computing Initiative

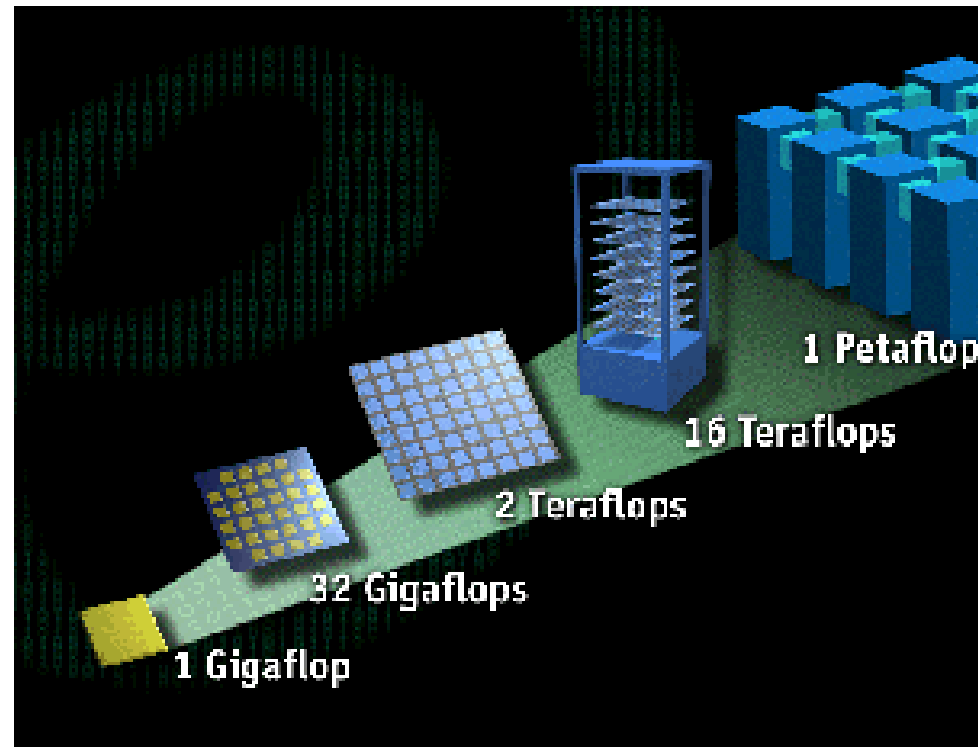
<http://www.llnl.gov/asci/>

- 1996 comprehensive testban on nuclear weapons signed;
- shift from nuclear test-based methods to computational-based methods of ensuring the safety, reliability, and performance of nuclear weapons stockpile
- create predictive simulation and virtual prototyping capabilities based on advanced weapon codes
- accelerate the development of high-performance computing far beyond what might be achieved in the absence of a focused initiative.

# ASCI (cont.)



# CMOS Petaflop/s Solution

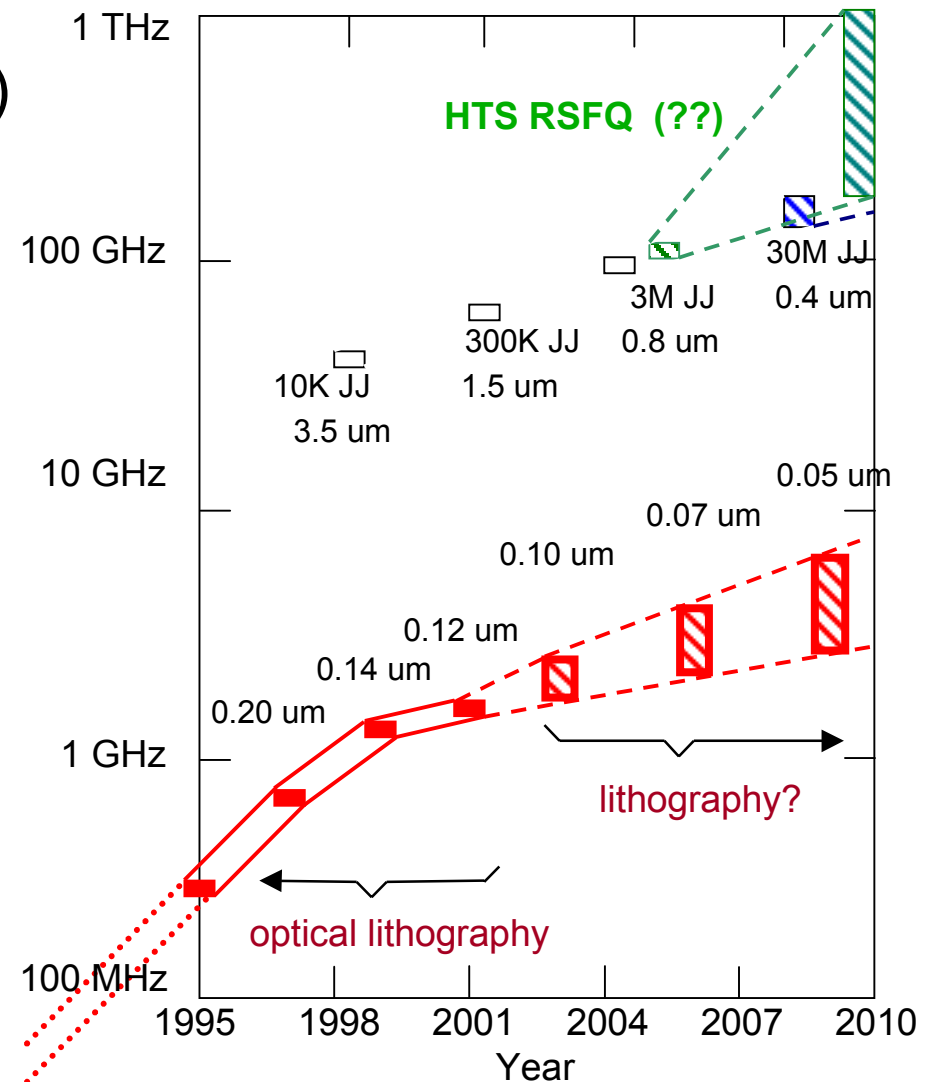
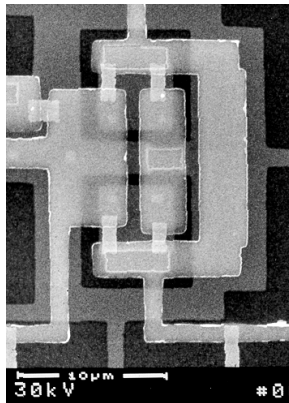


- IBM's Blue Gene
- 64,000 32 Gflop/s PIM chips
- Sustain  $O(10^7)$  ops/cycle to avoid Amdahl bottleneck

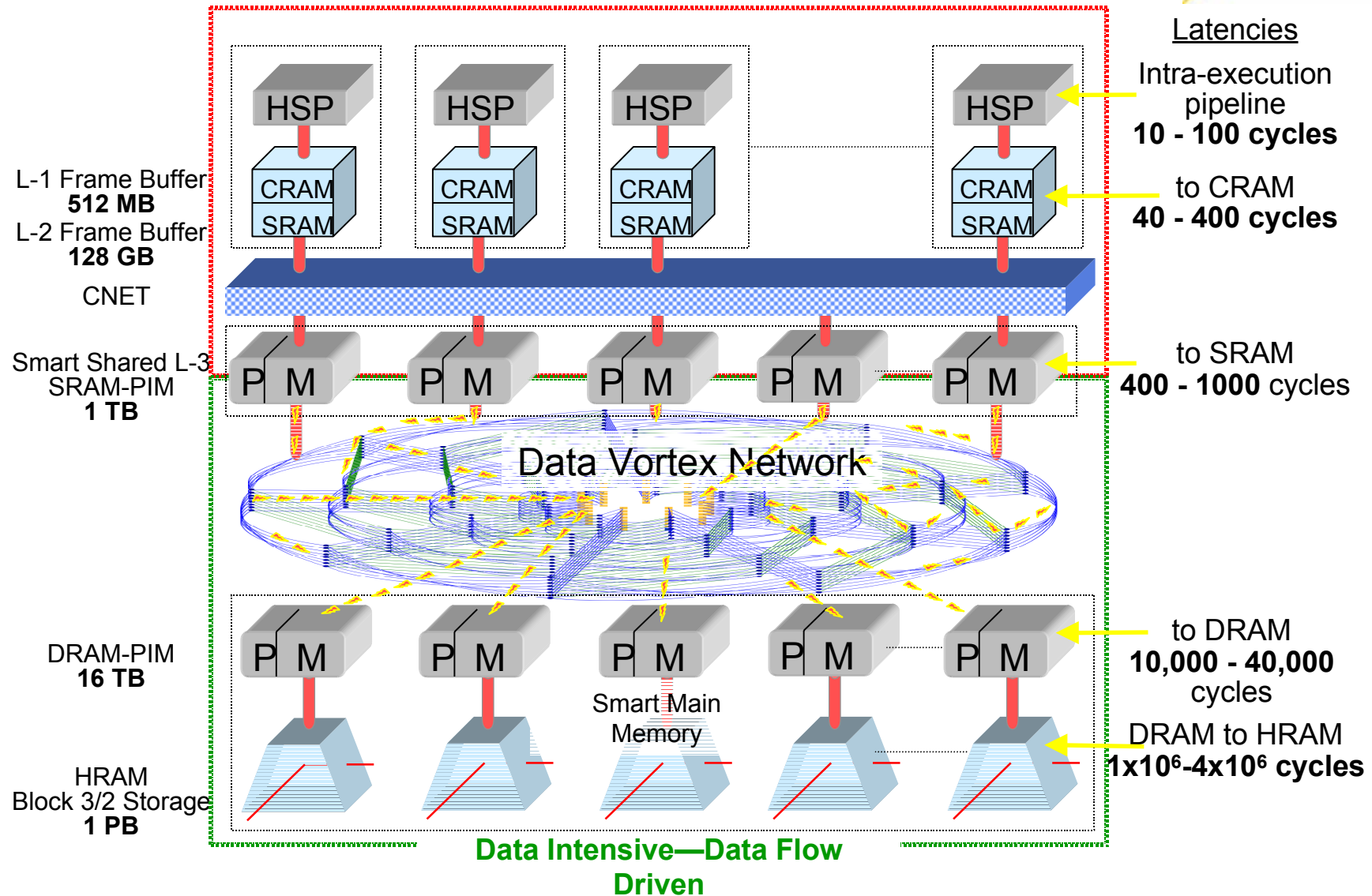
# An Alternate Technology?



- Single Flux Quantum (SFQ)
- Operates at 4 Kelvin



# Hybrid Technology, Multithreaded Architecture

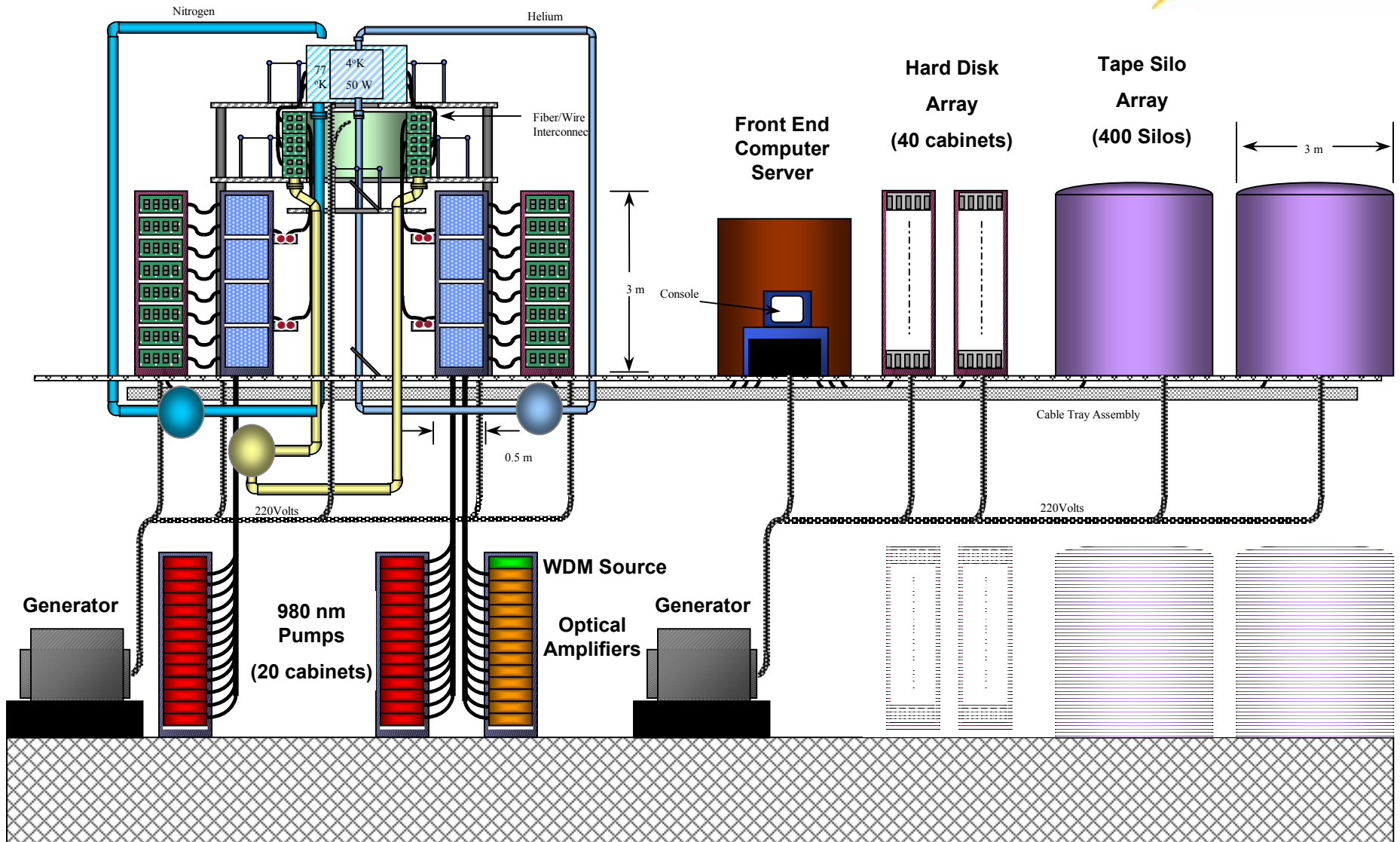


# HTMT Machine Room





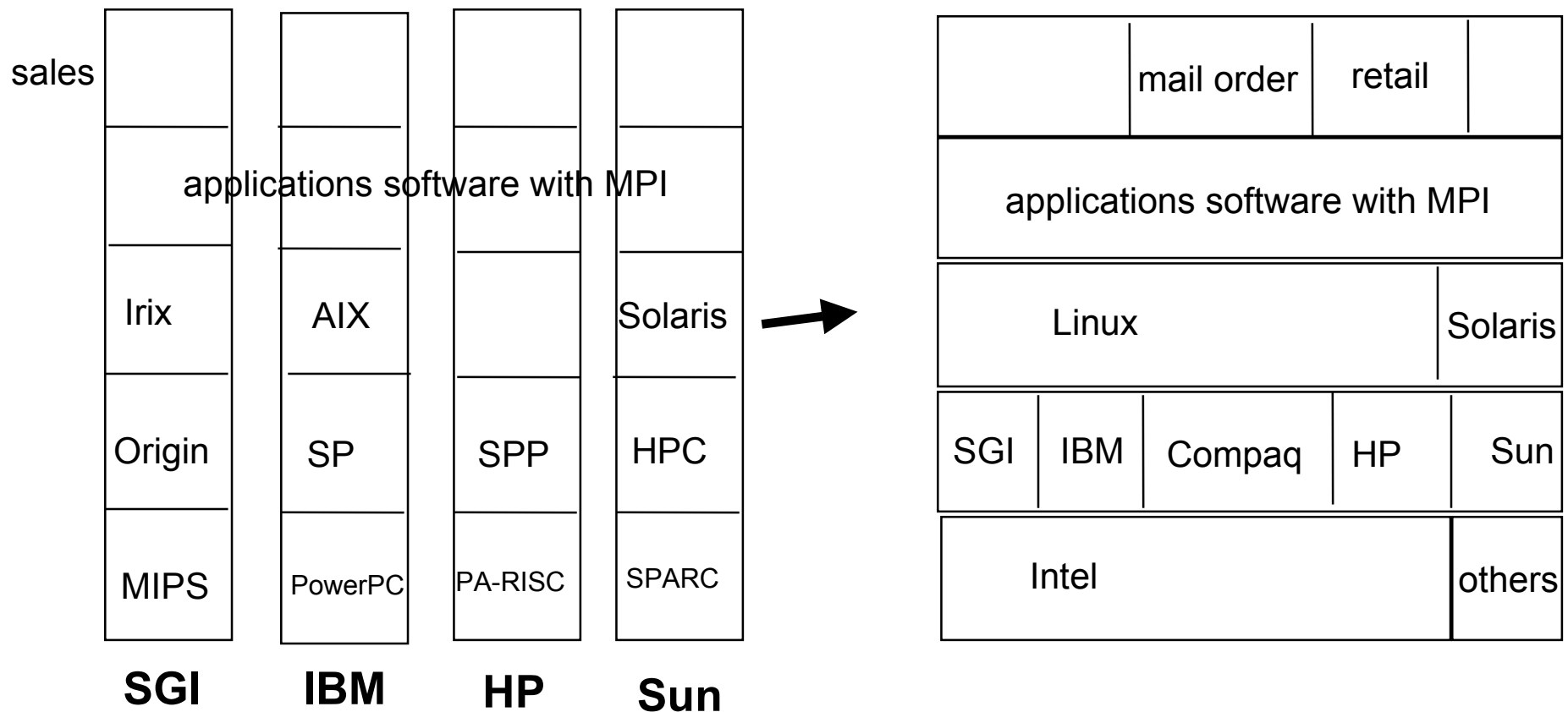
# HTMT Cross-Section



# 2000 - 2005: Market Issues



From vertical to horizontal companies - the **Compaq model of High Performance Computing**





# Until 2010: Market Issues

---



**Compaq's acquisition of DEC was just the first step. DEC transformed from vertical to horizontal in less than one year.**

**Business transition will be more fundamental than previous technology transition.**

**Tremendous impact on HPC community - no more business as usual (e.g. how do we procure machines)**

**Extremely difficult to pick winner**

**Tumultuous transition may make it difficult for boutique companies such as Cray, Inc. to survive**

# Contributions of Beowulf

---



- An experiment in parallel computing systems
- Established vision low cost high end computing
- Demonstrated effectiveness of PC clusters for some (not all) classes of applications
- Provided networking software
- Provided cluster management tools
- Conveyed findings to broad community
- Tutorials and the book
- Provided design standard to rally community!
- Standards beget: books, trained people, software ... virtuous cycle

Adapted from Gordon Bell, presentation at Salishan 2000

# Linus's Law: Linux everywhere

---



- Software is or should be free
- All source code is “open”
- Everyone is a tester
- Everything proceeds a lot faster when everyone works on one code
- Anyone can support and market the code for any price
- Zero cost software attracts users!
- All the developers write lots of code

# Open Source will change the rules!

---



- Stage 1: (40s and 50s): every computer different, every program unique
- Stage 2: (60s and 70s): software is unbundled from hardware, commercial software companies arise
- Stage 3: (80s and 90s): mass market computers and mass market software, the notions of software copyright and privacy are born
- Stage 4: (2000 and beyond): software migrate to the WWW, OSS communities provide high quality software, OSS takes over generic software

# Commercially Integrated Clusters are Already Happening

---



- Forecast Systems Lab procurement (Prime contractor is High Performance Technologies Inc., subcontractor is Compaq)
- Los Lobos Cluster (IBM with University of New Mexico)

## Linux Supercomputer Howls

by [John Gartner](#) and [Mi](#)

3:00 a.m. Mar. 22, 2000 PST

The University of New Mexico and IBM are teaming up to build the world's fastest Linux-based supercomputer.

Named "LosLobos", the new supercomputer is scheduled to be fully operational by the summer.

---

### TECHNOLOGY **Also:**

[SGI Sells a Piece of Its Heart](#)  
[Caldera Set to Ride Linux Wave](#)

News from the [Linux](#) front  
Read more [Technology](#) news

---

*ponsored by  
est. The  
dwidth to  
nge everything.*

*ay's Headlines  
1 p.m. Mar. 22,  
0 PST*

[JX](#)  
[ercomputer](#)  
[vls](#)

[abit, Don't Run](#)

[... in New](#)

LosLobos is a departure from the traditional supercomputer set-up. It's built from 256 IBM Netfinity servers.

The Netfinity servers are linked together using special clustering software and high-speed networking hardware, which causes the separate units to act as one computer, delivering a processing speed of 375 gigaflops, or 375 billion operations per second.

# Until 2010: New Technology

---



The software challenge: overcoming the **MPI barrier**

- MPI created finally a standard for applications development in the HPC community
- standards are always a barrier to further development
- the MPI standard is a least common denominator building on mid 80ies technology

# Enablers of pervasive technologies

---



- General accessibility through intuitive interfaces
- A supporting infrastructure, perceived valuable, based on enduring standards
- MOSAIC browser and World Wide Web are enablers of global information infrastructure

Source: Joel Birnbaum, HP, Lecture at APS Centennial, Atlanta, 1999



# Information appliances



- Are characterized by what they do
- Hide their own complexity
- Conform to a mental model of usage
- Are consistent and predictable
- Can be tailored
- Need not be portable



Source: Joel Birnbaum, HP, Lecture at APS Centennial, Atlanta, 1999

# IP On Everything



CS267, UC Berkeley

# In the 2010s: Pervasive Computational Modeling

---



Commodity consumer products

Example:

MOTOROLA, Pager Division, Boynton Beach, Florida

Applications: Radioss/Parallel Solids

ABAQUS Standard/Explicit

Alias - Render Industrial Designs

EFMASS, MDS, from H.P., MCSPICE

System: 8 CPU POWER CHALLENGE

2 GB Memory, 40GB Disk

Problem: Pager Case

- Battery Containment
- Electronics Integrity
- Display Life



# Towards Ubiquitous Computational Modeling

---



1985	1990	1995
specialized hardware Cray X-MP	specialized hardware Cray Y-MP	commodity hardware POWER CHALLENGE XL
nuclear weapons lab.	industrial company unique control resource	industrial company decentralized divisional resource
unique multimillion \$ product (weapons impact)	expensive consumer product \$10K (car crash)	mass consumer product \$1.99 (pager/cellular phone)

# Overview

---



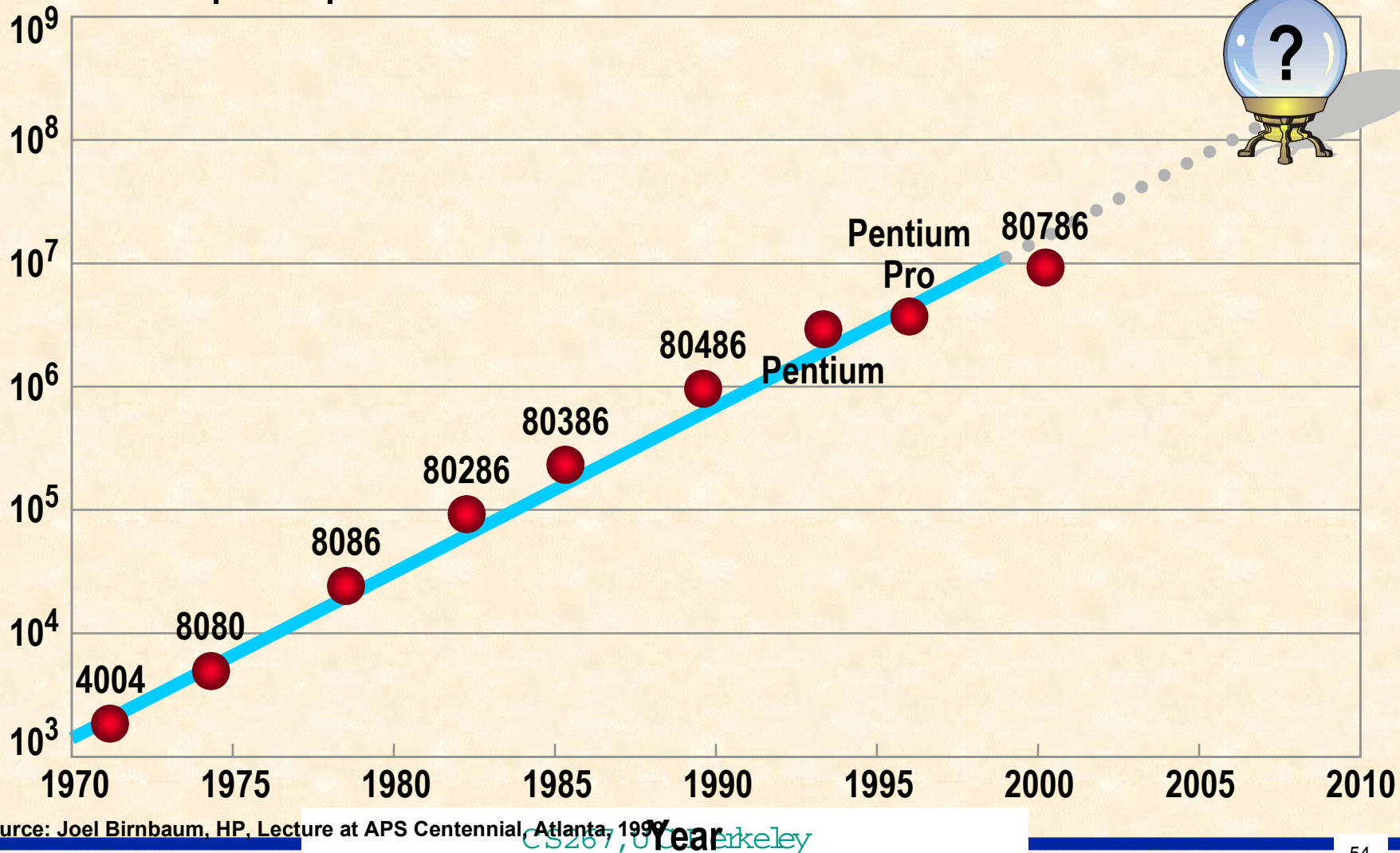
- Retrospective: changes in the 1990s
- Extrapolation to the near future up to 2010
- **The end of Moore's Law in about 2020**
- Beyond 2025



# Moore's Law



Transistors per chip



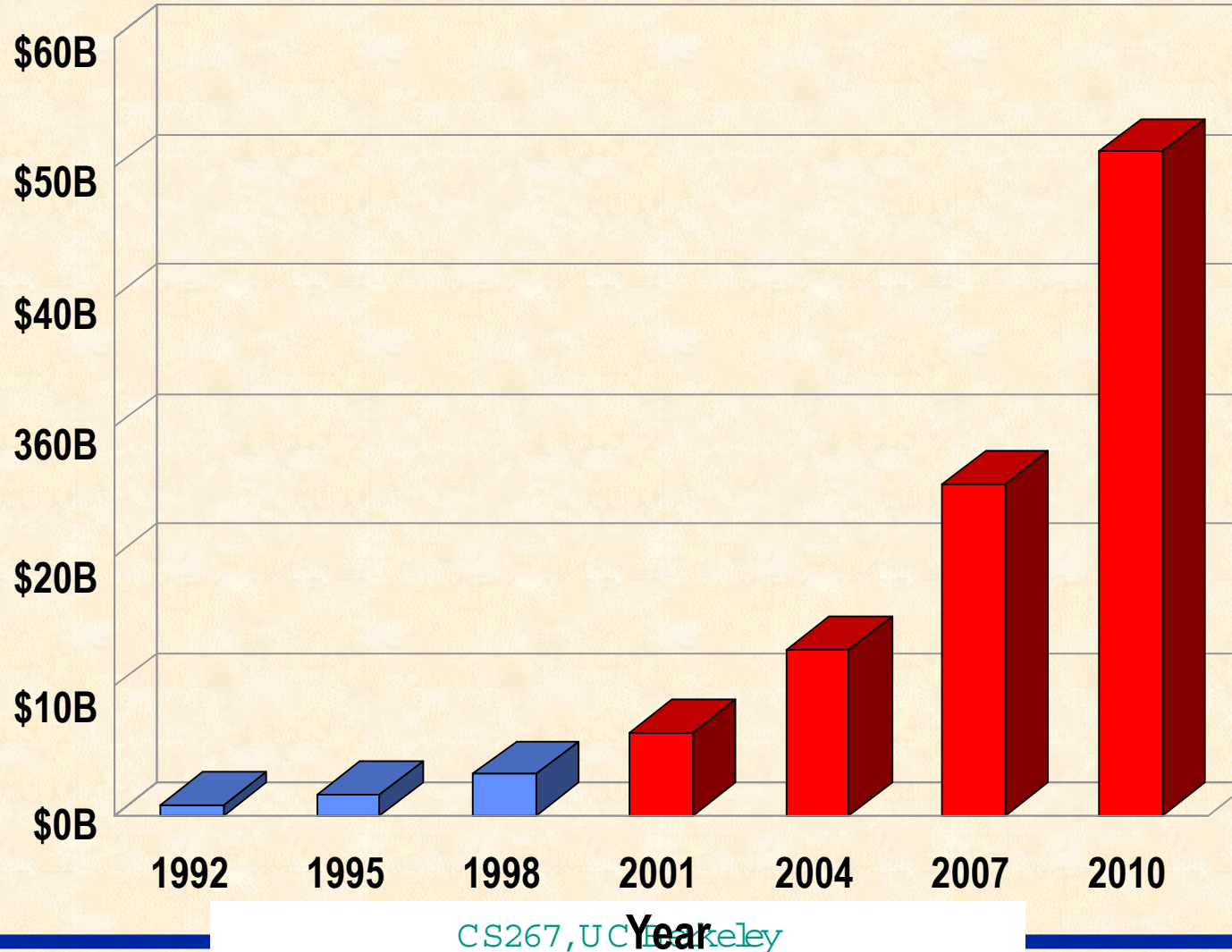
Source: Joel Birnbaum, HP, Lecture at APS Centennial, Atlanta, 1997

CS267, UC Berkeley

# Moore's Second Law

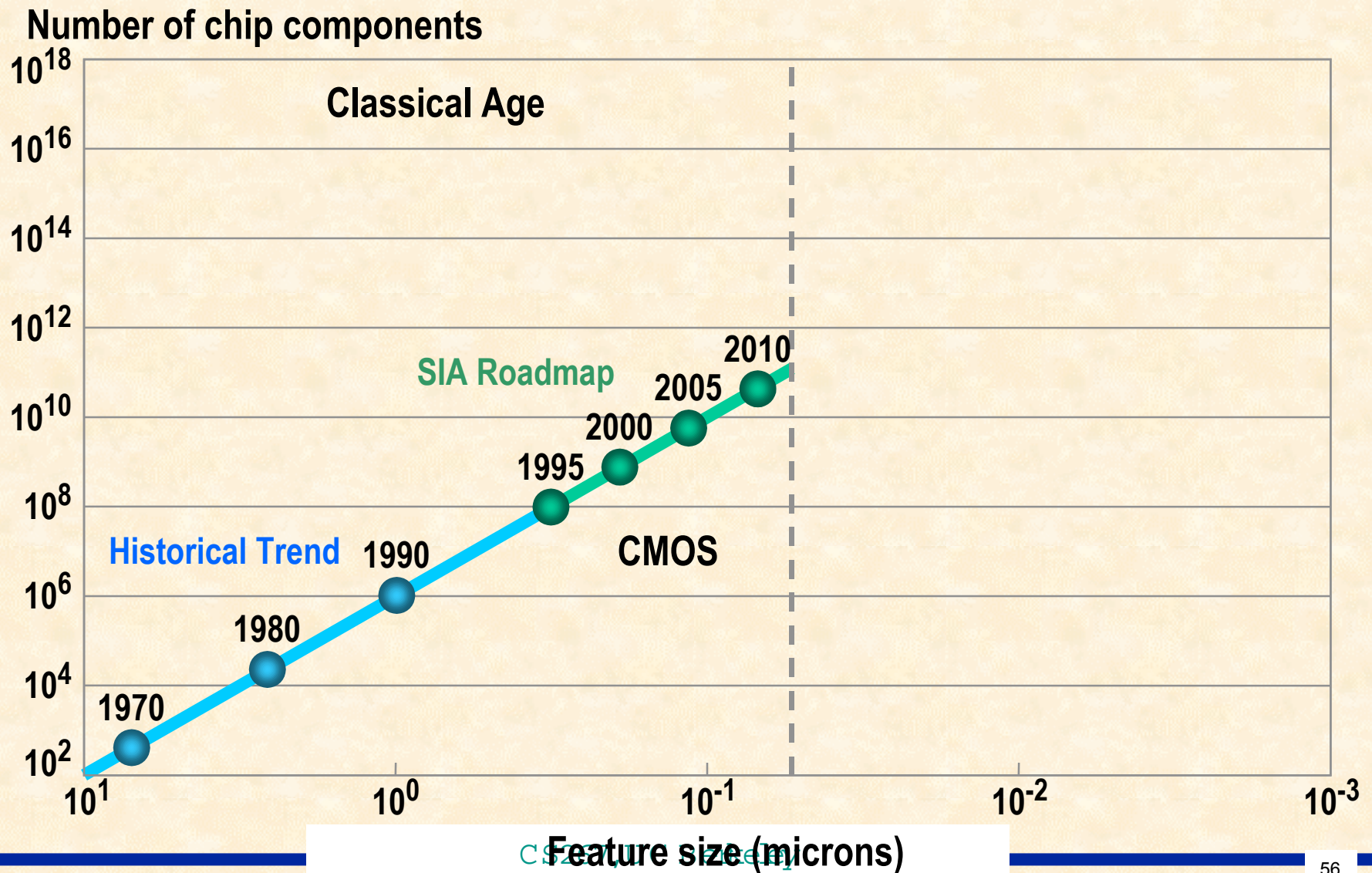


Cost of Fab



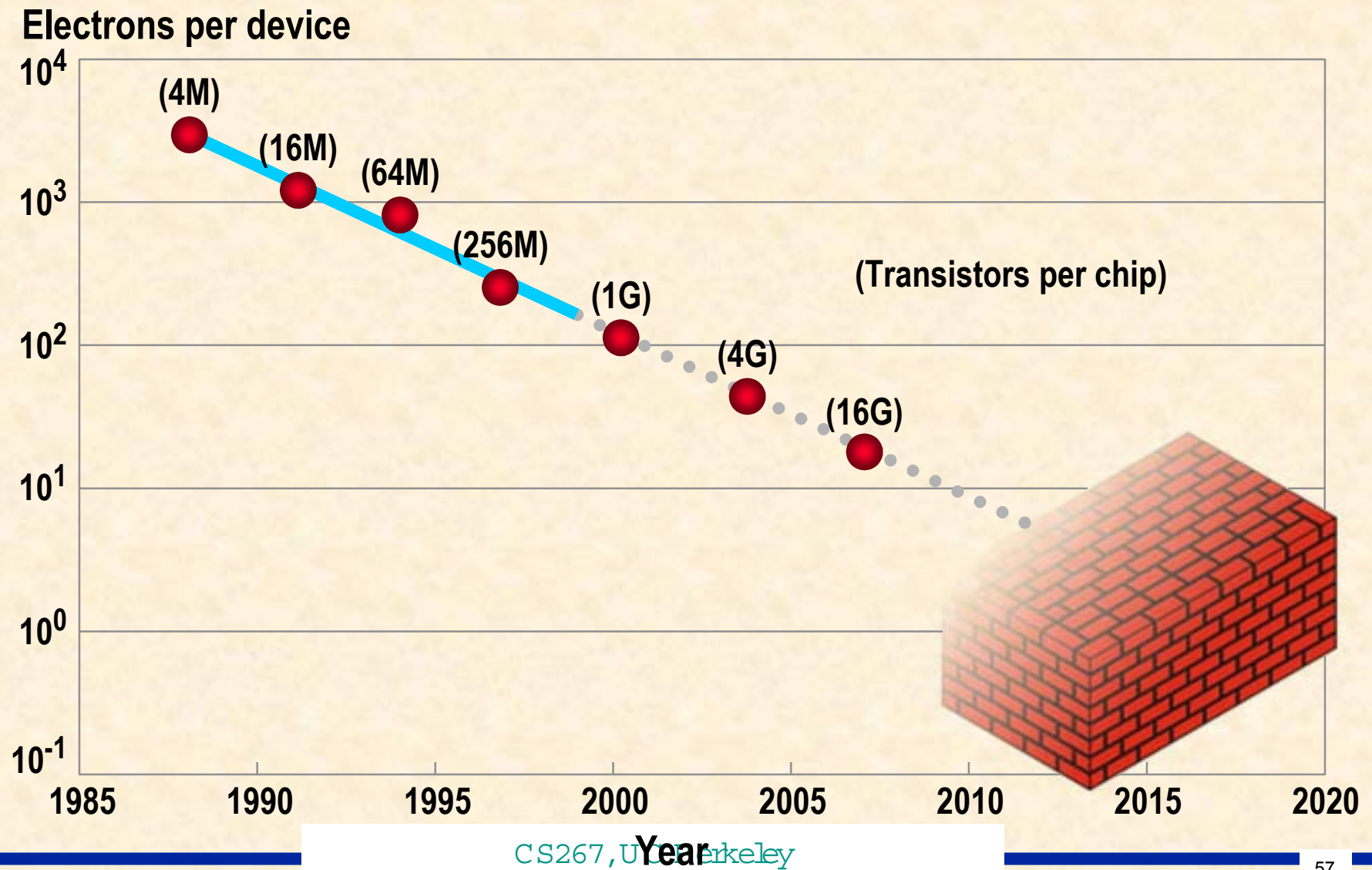
CS267, UC Berkeley

# Scaling of electronic devices

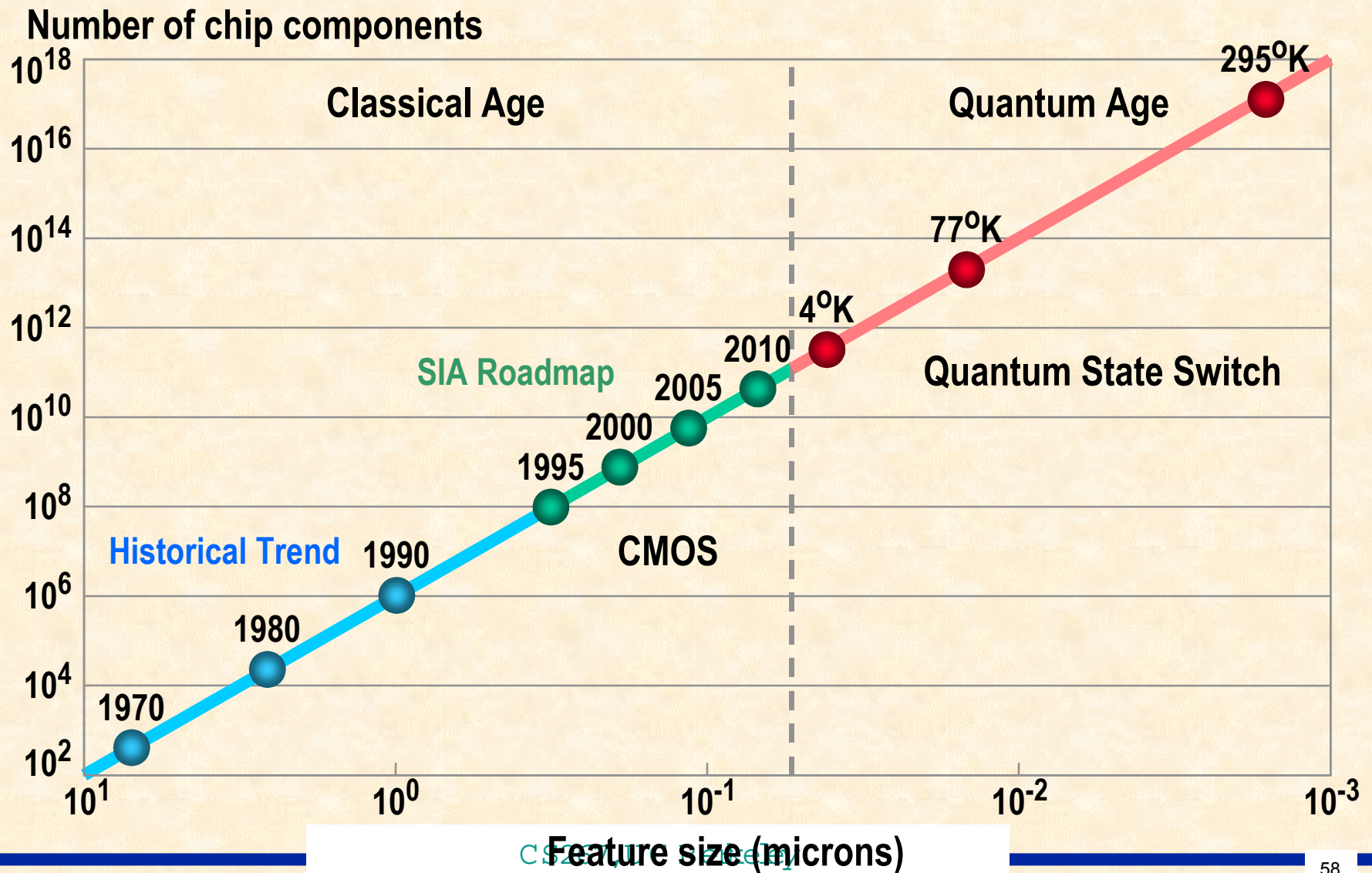




# Vanishing electrons



# Scaling of electronic devices



# Computation limit for nonreversible logic

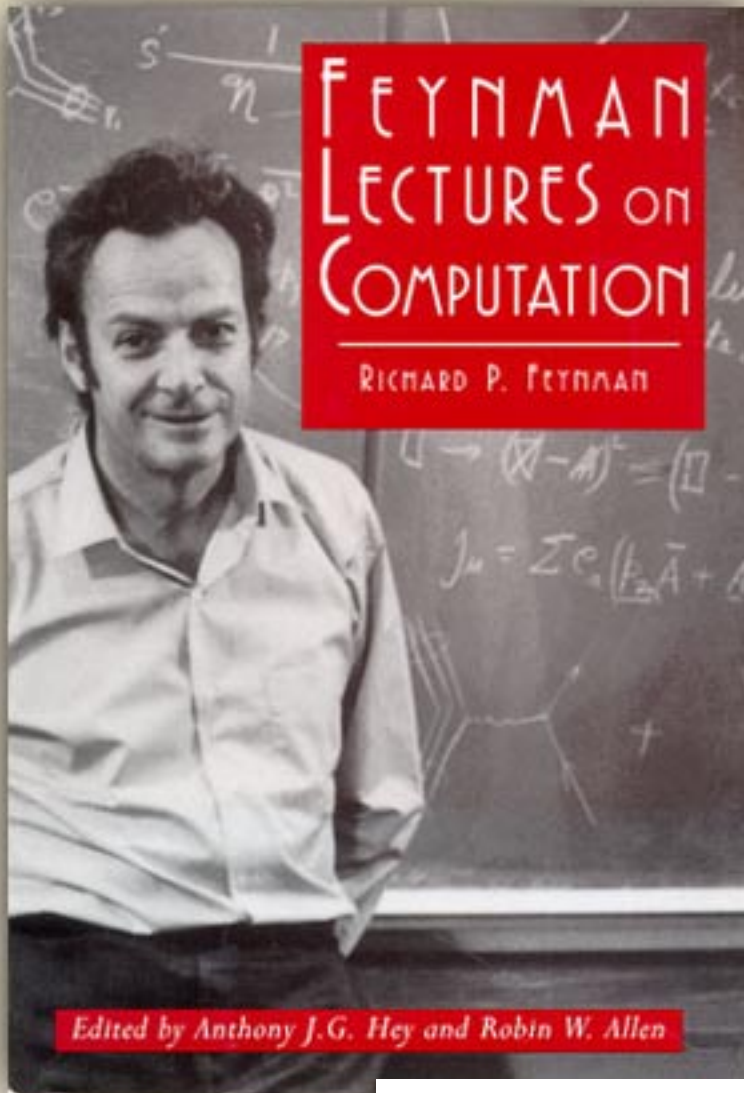


- Assume a power dissipation of 1W at room temperature

How many bit operation/second can be performed by a nonreversible computer executing Boolean logic?

$$\nu = P/kT \ln(2) = 3.5 \times 10^{20} \text{ bit ops/sec}$$

# Power cost of information transfer?



$$P = nk_B T \frac{d}{c} \nu^2$$

$P$  = power

$k_B$  = Boltzman constant

$T$  = temperature

$d$  = transmission distance

$c$  = speed of light

$\nu$  = operating frequency

$n$  = number of parallel operations



# Rate of nonreversible information transfer

---



- Assume a power dissipation of 1W and a volume of 1cm<sup>3</sup>

How many bits/second can be transferred?

$$v = \sqrt{\frac{cP}{k_B T d}} = 10^{18} \text{ ops/sec}$$

This is roughly the equivalent of 10<sup>9</sup> Pentiums!

# Other possibilities?

---



## Molecular nanomechanics:

- DNA, mechanical, chemical, biological

## Quantum cellular automata:

- Arrays of quantum dots

## Molecular nanoelectronics:

- Chemically-synthesized circuits

# Will history repeat itself?



	1939	1999
<b>Technology engine</b>	<b>Vacuum tube</b>	<b>CMOS FET</b>
<b>Disruptive technology</b>	<b>Solid state switch</b>	<b>Quantum state switch?</b>
<b>Fundamental research</b>	<b>Purity of materials</b>	<b>Size &amp; shape of materials</b>
<b>Impact</b>	<b>Demise of vacuum tubes</b>	<b>Demise of semiconductors</b>

# Thinking about 2025

---



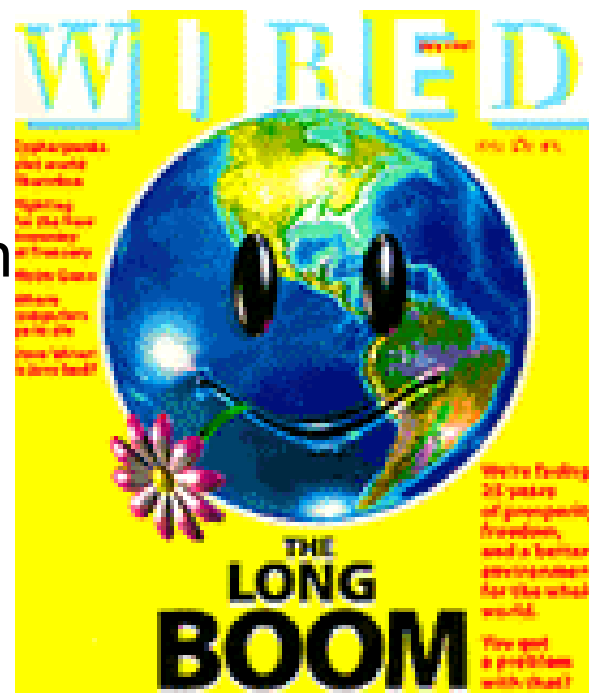
- Extrapolation
- “Reading the Clearing” (Denning)
- Scenario planning
- Science Fiction and Wishful Thinking



# Extrapolation: The Long Boom



- Peter Schwartz and Peter Leyden, Wired, July 1997
- global economic boom of unprecedented scale
- continued sustained economic growth
- managing ecological problems
- globalization and openness
- five waves of technology (computers, telecommunication, biotech, nanotechnology, alternative energy)



# Reading the Clearing: J. Coates, The Highly Probable Future c2025



- 8.4 B, English speaking, personally tagged & identified, prosthetic assisted and/or mutant, tense people who have access & control of their medical records
- Everything will be smart, responsive to environment.
  - Sensing of everything... challenge for science & engineering!
  - Fast broadband network
  - Smart appliances & AI
  - Tele-all: shop, vote, meet, work, etc.
  - Robots do everything, *but there may be conflict with labor...*
- A “managed”, physical and man-made world
  - Reliable weather reports
  - *“Many natural disasters e.g. floods, earthquakes, will be mitigated, controlled or prevented”*
- No surprises. We can see 10 years, but not 20!

Source: Gordon Bell and J. Coates, Futurist, Vol. 84, 1994

# Scenario Planning: Air Force 2025

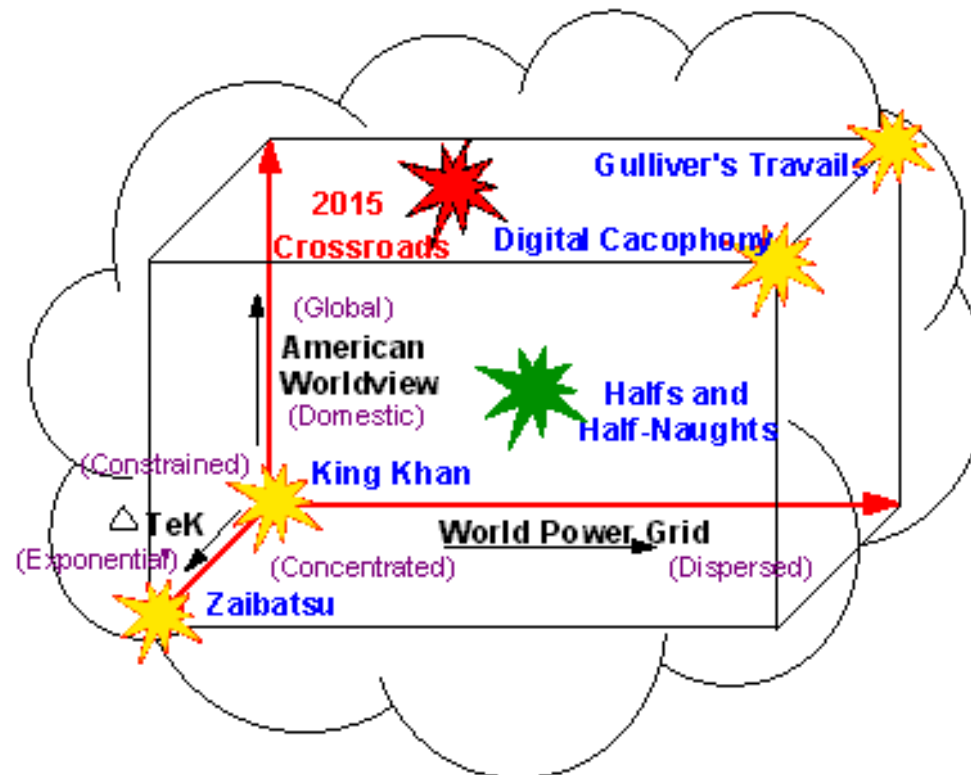


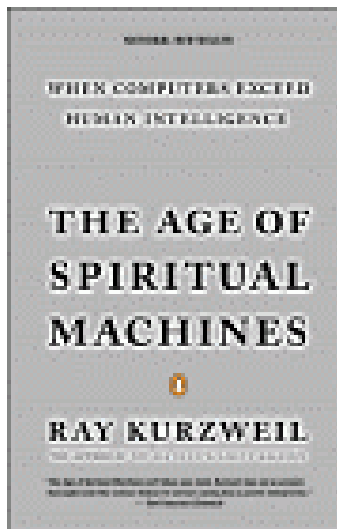
Figure ES-1. Strategic Planning Space

# Science Fiction and Wishful Thinking

---



- R. Kurzweil, The Art of Spiritual Machines
- Bill Joy, Why the Future Does Not Need Us, Wired March 2000



WIRED

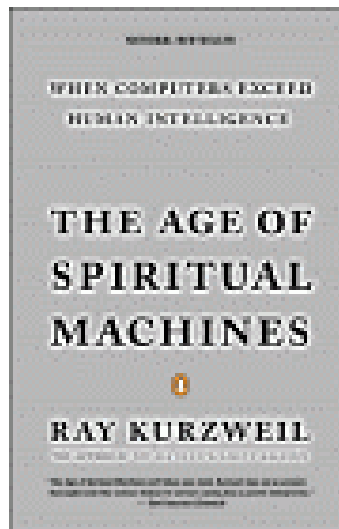


Why the Future  
Doesn't Need Us  
By Bill Joy

# Science Fiction and Wishful Thinking



- R. Kurzweil, The Art of Spiritual Machines
- Bill Joy, Why the Future Does Not Need Us, Wired March 2000



WIRED



Why the Future  
Doesn't Need Us  
By Bill Joy